



NONLINEAR PREDICTION AND COMPLEXITY OF ALPHA EEG ACTIVITY

MICHAEL E. BRANDT*

*Neurosignal Analysis Laboratory,
University of Texas Medical School, Houston, Texas, USA*

AHMET ADEMOĞLU

Biomedical Engineering Institute, Boğaziçi University, Istanbul, Turkey

WALTER S. PRITCHARD

*Psychophysiology Laboratory, R.J. Reynolds Tobacco Company,
Winston-Salem, North Carolina, USA*

Received December 22, 1998; Revised July 9, 1999

Two prediction techniques were used to investigate the dynamical complexity of the alpha EEG; a nonlinear method using the K-nearest-neighbor local linear (KNNLL) approximation, and one based on global linear autoregressive (AR) modeling. Generally, KNNLL has more ability to predict nonlinearity in a chaotic time series under moderately noisy conditions as demonstrated by using increasingly noisy realizations of the Hénon (a low-dimensional chaotic) and Mackey–Glass (a high-dimensional chaotic) maps. However, at higher noise levels KNNLL performs no better than AR prediction. For linear stochastic time series, such as a sine wave with added Gaussian noise, prediction using KNNLL is no better than AR even at very low signal-to-noise ratios. Both prediction techniques were applied to resting EEGs (O2 scalp recording site, 10–20 EEG system) from ten normal adult subjects under eyes-closed and eyes-open conditions. In all recordings tested, KNNLL did not yield a lower root mean squared error (RMSE) than AR prediction. This result more closely resembles that obtained for noisy sine waves as opposed to chaotic time series with added noise. This lends further support to the notion that these EEG signals are linear-stochastic in nature. However, the possibility that some EEG signals, particularly those with high prediction errors produced by a noisy nonlinear system cannot be ruled out in this study.

1. Introduction

Over the past decade, there have been numerous studies of the dynamical properties of the EEG based on measure-theoretic approaches such as correlation dimension [Babloyantz *et al.*, 1985; Mayer-Kress & Layne, 1987; Röschke & Başar, 1989; Rapp *et al.*, 1989; Pritchard & Duke, 1992a, 1992b; Pritchard *et al.*, 1995; Theiler & Rapp, 1996],

mutual information [Palus, 1993; Palus *et al.*, 1993] and Lyapunov exponents [Iasemidis *et al.*, 1990; Principe & Lo, 1991; Molinari & Dumermuth, 1992; Wallenstein & Nash, 1992]. The success and reliability of these measures depend on the validity of the underlying assumption of stationarity in appropriately long data segments, which is difficult to attain in EEG recordings. Additionally, specific choices of parameters chosen for analysis, use of a

*Address for correspondence: Department of Psychiatry & Behavioral Sciences, University of Texas-Houston Medical School, 6431 Fannin Drive, Room 5.202 Houston, Texas 77030-1501.

E-mail: mbrandt@ped1.med.uth.tmc.edu

particular control test for determinism, and choice of specific algorithm employed are crucial factors in attempting to uncover the fundamental nature of the underlying system which produces the measured EEG activity (see [Jansen, 1991], for a discussion of these issues). An alternative approach to determining the dynamical properties of the EEG is to use modeling and prediction techniques that may yield information about the signal within a relatively short interval of time during which it may be stationary.

Several recent studies have attempted to predict EEG activity [Blinowska & Malinowski, 1991; Pezard *et al.*, 1994; Hernandez *et al.*, 1995]. The main purpose of these studies was to characterize EEG produced by a presumed dynamical system which may be either linear or nonlinear in nature. Blinowska and Malinowski [1991] applied both linear autoregressive (AR) modeling and local nonlinear simplex methods [Sugihara & May, 1990] to predict the EEG. Their conclusion was that the EEG was not predictable for intervals longer than about 40 to 60 ms. They also pointed out that the nonlinear method used was not significantly better than the linear technique in terms of prediction performance. Poor predictability of EEG beyond several time samples was also confirmed by Parikh and Pratap [1991]. On the other hand, Kemp and Lopes da Silva [1991] showed a wide range of variability in EEG intra-individual prediction related to changes in physiological brain state. Casdagli [1992] concluded that even though there was little evidence for low-dimensional chaos in alpha EEG, this did not necessarily imply that the linear models were adequate to describe the time series. Since 1985, a majority of the published findings concerning EEG complexity in which a control test for determinism is employed actually do not support the general hypothesis that the EEG is produced by a deterministic chaotic process [Stam *et al.*, 1995]. This, however, does not necessarily rule out that it can be explained by nonlinear dynamics generally. Hernandez *et al.* [1995] applied both linear and nonlinear modeling to investigate predictability patterns in the EEG. They found that alpha EEG demonstrated three qualitative patterns: predictable for less than a few ms, moderately predictable (10 to 50 ms), and highly predictable (more than 50 ms). They concluded that nonlinear methods were more adequate than linear ones for EEG modeling and prediction.

In this study, eyes-closed and eyes-open scalp-recorded EEG activity from normal human adult subjects is analyzed using prediction techniques based on linear autoregressive (AR) modeling and a nonlinear technique referred to as K-nearest-neighbor local linear (KNNLL) approximation [Farmer & Sidorowich, 1988; Casdagli, 1989]. The major aim is to characterize the dynamical characteristics of this activity in terms of its linear or nonlinear dynamical properties and relate the results to those obtained for correlation dimension (D_2) using the Takens–Ellner (TE) algorithm [Takens, 1981, 1985; Ellner, 1988] in order to assess its dynamical complexity.

2. Methods

2.1. Approximation of the nonlinear map

Given a time series s_i , a state space trajectory is reconstructed using m -dimensional state vectors

$$\mathbf{x}_i = [s_i, s_{i+\tau}, \dots, s_{i+(m-1)\tau}], \quad (1)$$

where τ is the time delay and m is the embedding dimension [Takens, 1981]. For τ -step ahead prediction of s_i , a nonlinear map f can be defined on this trajectory as

$$f(\mathbf{x}_i) = s_{i+m\tau}, \quad (2)$$

where f can be approximated using local or global polynomial kernels [Farmer & Sidorowich, 1988; Casdagli, 1989; Stokbro & Umberger, 1992], neural networks [Weigend *et al.*, 1992], radial basis functions [Casdagli, 1989; Lowe & Webb, 1991] or wavelets [Cao *et al.*, 1995]. A local nonlinear approximation to f fits a hypersurface to the K-nearest-neighbors of \mathbf{x}_i . In the KNNLL approach, hyperplanes are used to approximate f . The time series s_i is divided into training and prediction data sets. For a vector \mathbf{x}_p in the training set, the corresponding K-nearest-neighbor vectors $\mathbf{x}_{p_1} \cdots \mathbf{x}_{p_K}$ in the training set are determined using a suitable norm (e.g. Euclidean), and a least-squares optimization is performed based on their τ -step ahead predictions

$$\begin{aligned} f(\mathbf{x}_{p_1}) &= s_{p_1+m\tau} \\ f(\mathbf{x}_{p_2}) &= s_{p_2+m\tau} \\ &\vdots \\ f(\mathbf{x}_{p_K}) &= s_{p_K+m\tau}. \end{aligned} \quad (3)$$

If f is to be approximated by a hyperplane with parameters \mathbf{L}_p in the vicinity of the vector \mathbf{x}_p , then Eq. (3) becomes

$$\begin{aligned} & \begin{bmatrix} 1 & s_{p_1} & s_{p_1+\tau} & \cdots & s_{p_1+(m-1)\tau} \\ 1 & s_{p_2} & s_{p_2+\tau} & \cdots & s_{p_2+(m-1)\tau} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & s_{p_K} & s_{p_K+\tau} & \cdots & s_{p_K+(m-1)\tau} \end{bmatrix} \begin{bmatrix} L_{p_0} \\ L_{p_1} \\ L_{p_2} \\ \vdots \\ L_{p_m} \end{bmatrix} \\ &= \begin{bmatrix} s_{p_1+m\tau} \\ s_{p_2+m\tau} \\ \vdots \\ s_{p_K+m\tau} \end{bmatrix}, \end{aligned} \quad (4)$$

where $\mathbf{L}_p = [L_{p_0} L_{p_1} L_{p_2} \cdots L_{p_m}]$. Least-squares optimization of Eq. (4) will yield a local linear approximation of f around the vector \mathbf{x}_p . The hyperplane parameter vector \mathbf{L}_p can be used for τ -step ahead prediction as

$$\hat{s}_{p+m\tau} = [1 \ \mathbf{x}_p] \mathbf{L}'_p. \quad (5)$$

Using all the training vectors (as neighbors) for optimization corresponds to AR modeling which can be expressed in terms of the data points s_i as

$$s_i = \sum_{k=1}^m a_k s_{i-k} + \varepsilon_i, \quad (6)$$

where the a_k 's are the AR coefficients, m is the model order and ε_i represents the unpredictable portion of the signal.

2.2. Prediction method

To perform prediction, the data is divided into training and prediction sets. Then the nonlinear approximation of the trajectory is computed from the vectors in the training set. The approximated map is used for predicting the future values in the prediction set. More specifically, given a time series with N_1 points for training and N_2 points for prediction, the number of embedding vectors are $N_1 + N_2 - (m-1)\tau$. The first $N_1 - (m-1)\tau - 1$ vectors are used to approximate the map. Starting from the $N_1 - (m-1)\tau$ th vector, the next $N_T = N_2 - M\tau + 1$ vectors are input to the map to perform N_T τ -step ahead predictions, where M is the maximum number of iterative prediction steps. For $p\tau$ -step prediction, the prediction routine is iterated p times for

each vector chosen from the prediction set. The relative mean squared error (RMSE) for $p\tau$ -step iterative prediction is calculated as:

$$\varepsilon(p) = \frac{\sum_{k=1}^{N_T} (s_{k+q} - \hat{s}_{k+q})^2}{N_T \sigma^2}, \quad \text{for } q = N_1 + (p-1)\tau, \quad (7)$$

where $\{\hat{s}_i\}$ and $\{s_i\}$ are the sets of predicted and real data points respectively and σ is the standard deviation of $\{s_i\}$.

2.3. Stationarity and complexity analysis

The Phillips–Perron (P–P) unit-root test [Phillips & Perron, 1988] is used to check the cointegration stationarity of the data segments used in the analyses. Such cointegration stationarity is an assumption of all AR models. Specifically, the P–P test is a test of the null hypothesis that $\beta = 0$ in the following AR(1) equation:

$$\Delta s_i = \mu + \beta s_{i-1} + \varepsilon_i, \quad (8)$$

with ε being a zero-mean, constant variance independent, identically distributed (IID) disturbance, and Δ the difference operator. The equation is estimated by ordinary least squares and then the t -statistic of β is corrected for serial correlation in ε_i . The order q of the serial correlation is determined using the Newey–West procedure, which is based (solely) on the number of observations N :

$$q = 4 \left(\frac{N}{100} \right)^{2/9}, \quad (9)$$

For the derivation of this equation and further technical details refer to [Newey, 1987]. The P–P test is applied to each entire data segment (1024 points) for all subjects. The 1% critical value for the data consisting of 1024 points is -3.4395 .

For each 1024-point EEG segment, the correlation dimension D_2 was estimated using the TE method. This method is conceptually similar to the more familiar Grassberger–Procaccia algorithm [Grassberger & Procaccia, 1983] but is based on a change of variable from distance (between reconstructed state-space vectors) to the negative log of distance. This change results in a distribution having a region of exponential behavior from which a maximum-likelihood estimate of D_2 can be made

using the basic properties of order statistics and the exponential distribution. For each embedding m , the embedding lag τ was set according to the geometry-based method of Rosenstein [Rosenstein *et al.*, 1994]. Like Rosenstein *et al.*, we have found this method to work very well with known simulated time series. This method is based upon the rate of expansion of the reconstructed state-space vectors from the main state-space diagonal. Using this lag, the state-space trajectory was reconstructed as a set of $N_v = N - (m - 1)\tau$ vectors of length m (N is the number of points in the time series). Singular-value decomposition was then employed to eliminate noise dimensions in the embedding. $N_v/2$ pairs of vectors were randomly chosen, and a set of $r(i)$ distances were computed for $i = 1, \dots, N_v/2$, using the Euclidean norm. Each $r(i)$ was then transformed to $\rho(i)$ by the formula $\rho(i) = -\ln[r(i)]$. The values of $\rho(i)$ were then sorted into ascending order, forming a set of order statistics drawn independently from the cumulative distribution function $F(\rho)$. The transformation and sorting permit a maximum-likelihood approach to D_2 estimation using methods derived from the statistical theory of reliability [Mann *et al.*, 1974], but otherwise is equivalent to Grassberger–Procaccia.

For the range of $[I_0 = (0.78)(N_v/2)] < i < [I_1 = (0.98)(N_v/2)]$, we found empirically that the sorted values of $\rho(i)$ had a probability distribution that was approximately a (truncated) exponential of density $F(\rho) = D_2 \exp(-D_2\rho)$, [Ellner, 1988]. Letting $M = I_1 - I_0$, the sum, $Y = y(1) + y(2) + \dots + y(M)$ was computed, where $y(i) = \rho(i + I_0) - \rho(I_0)$. This subtraction effectively takes into account the fact that all values of $r < r(I_0)$ are to be discarded. The maximum likelihood estimate is given as

$$D_2 = \frac{M}{Y + y(M)[N_v/2 - I_1]} \quad (10)$$

[Ellner, 1988]. The second term in the denominator of this expression accounts for values of $r < r(I_1)$ that fall outside the window of approximate exponential behavior. The sampling variability of this estimate was reduced by bootstrapping using the exact procedure outlined by Ellner [1988], namely repeated random sampling with replacement of $N_v/2$. We bootstrapped each estimate 64 times, each time choosing different pairings from the full set of reconstructed vectors.

The method of surrogate data analysis [Theiler *et al.*, 1992] is used to detect the nonlinearity in

a time series. There are several methods of generating a surrogate data set from an original time series. This can be achieved either by computing the Fourier components of the signal and randomizing the phases of these frequency components while keeping their amplitude the same or modeling the signal using a linear AR model and obtaining different realizations of the process by generating white noise and passing them through the AR filter. In the former case, it is also possible to generate a surrogate time series matching the amplitude distribution of the original signal. The essential aim of these methods is to generate a time series having similar power spectral properties as the original signal while removing all the nonlinear autocorrelations of the original time series. It is assumed that D_2 of a nonlinear time series will be significantly different from that of its surrogates. In the present study, for each EEG time series, 16 Gaussian amplitude-adjusted, phase-angle randomized surrogates were produced [Theiler *et al.*, 1992]; the Gaussian amplitude adjustment controls for the possibility of a static, nonlinear measurement transform of normally distributed data. The above D_2 estimation procedure was applied to each EEG surrogate, with a final surrogate D_2 estimate computed as the average across the 16 surrogate D_2 estimates.

2.4. Simulated data

To test their validity as predictors, both the linear AR and nonlinear KNNLL methods were applied to simulated signals. The data consisted of chaotic time series produced from the Hénon (low-dimensional) and Mackey–Glass (high-dimensional) systems and a sinusoid with peak amplitude of 1.0, frequency of 10 Hz, and sampling rate of 128 Hz. The Hénon map is given as

$$x_{n+1} = 1 - ax_n^2 + bx_{n-1}. \quad (11)$$

The parameters were chosen as $a = 1.4$ and $b = 0.3$ in which the map operates in the chaotic regime. Initial conditions were $x(0) = x(-1) = 0.01$. The Mackey–Glass system is given by

$$\frac{dx(t)}{dt} = -0.1x(t) + \frac{0.2x(t-T)}{1 + x(t-T)^{10}}. \quad (12)$$

We fixed the parameter $T = 100$ and the initial condition as $x(t) = 0.9$ for $0 \leq t \leq T$. The sampling rate was 6. The first 1000 points generated from the chaotic maps were always discarded to reduce the transient effects.

Four different levels of noise were added to each of these three signals and prediction was performed. The standard deviations of the white Gaussian pseudorandom sequences were chosen as 25%, 50%, 75% and 100% of the standard deviation of each simulated (original time series). The number of data points in each signal was 1024. The first half of the data samples was used for training and the second half for prediction. The optimal m values for the Hénon, Mackey–Glass and sine data were 2, 12 and 3, respectively. These values were determined by varying the m values from 2 to 20 and selecting the one yielding the lowest root mean square prediction error. The τ (lag) value for the Hénon map was 1 and those for the Mackey–Glass and sine data were both 3. This was the lag time corresponding to the autocorrelation function dropping to its $(1/e)$ th value.

2.5. EEG data collection

In order to apply the above prediction methods to the EEG data, four 8 s, artifact-free alpha EEG segments from a 120 s recording period were selected from ten neurologically normal adult subjects under both eyes-closed and eyes-open conditions. The three segments having the highest alpha-band power were then selected for analysis (for the eyes-open condition, three with the lowest alpha-band power were selected). The electrode placements were made from the O1 and O2 scalp loci (10–20 EEG system) referenced to nasion, with a forehead ground. In the following, data analysis was carried out on the O2 recordings only. Tin electrodes with impedances less than 3 k Ω were used. Bandpass filter settings were 0.3 to 30 Hz, with -12 dB/octave rolloff, and the sampling rate was 128 Hz. The number of training (N_1) and prediction (N_2) data points were both fixed as 512 (4 s). These values were

found to correspond to stationary EEG data segments. For prediction, the lag (τ) was set to 3 in all sets. Both the lag time and optimal value for m were selected using the same procedure as described above for the simulated data sets.

3. Results

Table 1 lists the minimum RMSE for one-step ahead prediction of simulated signals (Hénon, Mackey–Glass, sine wave) with 25, 50, 75 and 100 percent noise levels added. For the KNNLL method, prediction performance was tested for k values varying from 1 to 511, and the minimum RMSE for one of the k values in that range is listed in Table 1. Using all 512 training vectors yields the AR RMSE as shown. Even with a small number of data points (1024), the technique demonstrates its sensitivity to nonlinear data. As can be observed from Table 1, even with 75% noise added to the signal, the KNNLL method is at least as effective as the AR method in prediction of low- and high-dimensional chaotic signals based on the one-step ahead RMSE. Figures 1–3 show the prediction results for all three simulated mappings as a function of the number of nearest neighbors (k). For the nonlinear mappings (Hénon and Mackey–Glass) with no or low-amplitude additive noise, there is a distinctive minimum RMSE for a particular k value well below the maximum tested (all the vectors corresponding to the AR method), and AR prediction yields the lowest RMSE only when 100% noise is added to each of the two maps. In Fig. 2 a seemingly paradoxical result is observed: prediction of the Mackey–Glass system with 25% noise level added appears better than with no noise added! First, this seeming difference may not be statistically significant, especially given that this is a high-dimensional

Table 1. Minimum RMSE for one-step prediction of simulated signals with indicated noise levels.

Noise Level	Hénon		Mackey–Glass (100)		Sine Wave (10 Hz)	
	KNNLL	AR	KNNLL	AR	KNNLL	AR
25%	0.33	0.87	0.47	0.81	–	0.07
50%	0.63	0.91	0.72	0.85	–	0.23
75%	0.88	0.96	0.88	0.91	–	0.41
100%	–	0.88	–	0.91	–	0.56

“–”: Corresponds to cases where KNNLL method does not yield a lower error than the AR method.

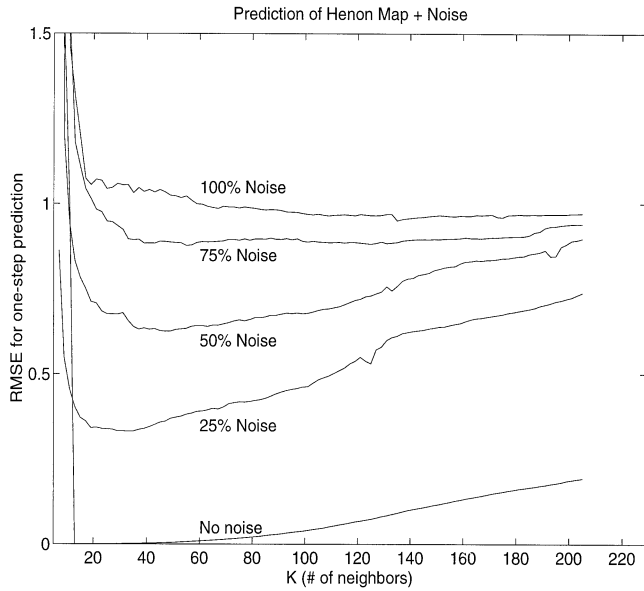


Fig. 1. Prediction results for the Hénon map with additive noise of 25 to 100%.

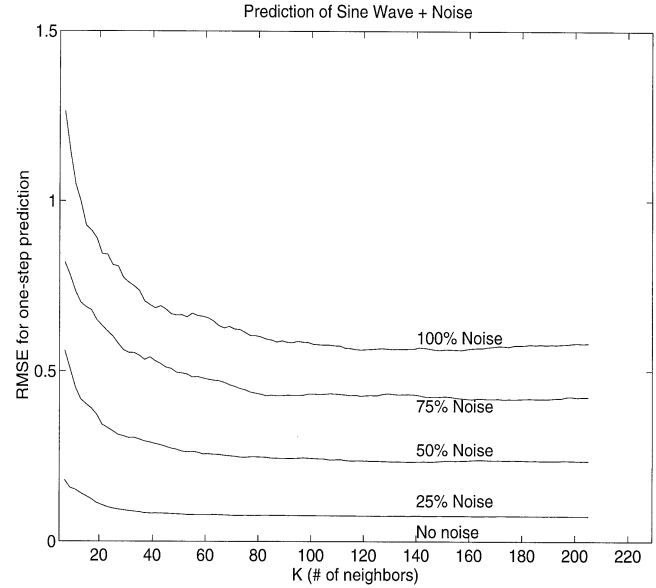


Fig. 3. Prediction of the sine wave with additive noise of 25 to 100%.

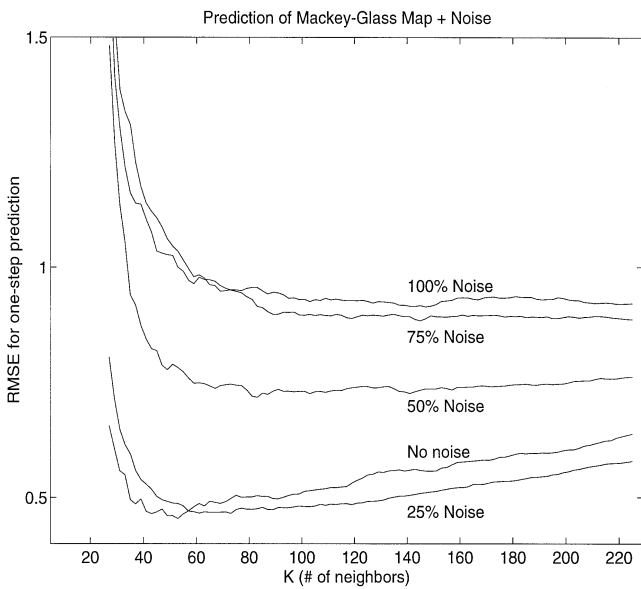


Fig. 2. Prediction of the Mackey–Glass map with additive noise of 25 to 100%.

system to begin with. Second, if this difference happens to be significant, it may be because the addition of a moderate amount of noise can in some systems increase the ability of both linear and nonlinear prediction methods to generalize the training set and thus actually enhance the ability to predict the test set. This is a well-known phenomenon in training–testing paradigms using neural networks (of which the KNNLL prediction method employed

here is related to). For the sine wave data, KNNLL prediction is not effective at any noise level (that is, the minimum RMSE occurs when all nearest neighbors are used, corresponding to AR prediction).

The prediction results for the EEG data are given in Table 2. In all cases, AR prediction yields the lowest RMSE. Stated alternatively, nonlinear prediction using KNNLL is less effective than linear prediction using AR modeling for all EEG segments included in this study. Figure 4 shows the eyes-open and eyes-closed EEG and prediction results (RMSE versus k) for subject number 5 who demonstrated the lowest AR prediction (for the eyes-closed data). Figure 5 shows the data for subject number 2 who had the largest eyes-closed AR prediction RMSE. Repeated measure analysis of variance (ANOVA) revealed that the RMSE for the eyes-open EEGs were significantly greater than for the eyes-closed EEGs ($F = 34.7, p < 0.0002, df = 1$).

In Table 3, P–P statistics, D_2 and the δD_2 (surrogate D_2 minus the original D_2 , reflecting nonlinearity in the EEG) are given for each 1024-point EEG segment. The P–P values for both eyes-closed and eyes-open EEG are well below the 1% critical value indicating that the data are stationary with respect to cointegration. P–P scores and D_2 of the eyes-open EEGs are significantly greater than those of the eyes-closed EEGs (for P–P, $F = 26, p < 0.0006, df = 1$; for D_2 , $F = 9.4, p < 0.01, df = 1$). Also, in contrast to the eyes-closed, is

Table 2. Minimum RMSE for one-step AR prediction of EEG series (KNNLL prediction yields higher RMSE in all these cases).

Subject #	Eyes Closed	Eyes Open
1	0.73	0.90
	0.90	0.96
	0.78	0.93
2	0.37	0.68
	0.30	0.82
	0.27	0.96
3	0.45	0.69
	0.37	0.71
	0.35	0.76
4	0.39	0.89
	0.29	0.78
	0.19	0.53
5	0.96	1.02
	1.02	0.76
	0.72	0.78
6	0.53	0.98
	0.31	0.87
	0.43	0.72
7	0.31	0.66
	0.41	0.78
	0.49	0.73
8	0.34	0.62
	0.38	0.60
	0.36	0.75
9	0.28	0.86
	0.29	0.82
	0.42	0.97
10	0.28	0.61
	0.27	0.89
	0.49	0.75
Mean	0.45	0.70

the fact that D_2 increases with decreasing P-P. There was no significant difference between the eyes-open and eyes-closed δD_2 .

4. Discussion

If we compare the results shown in Figs. 4 and 5 with those of the simulated data in Figs. 1–3 and Tables 1 and 2, we may reasonably conclude that the EEGs in this study must either be produced by a nonlinear system with dimensionality higher than that of the Mackey–Glass system (greater than 10), or by a linear stochastic system. Based on the D_2 's

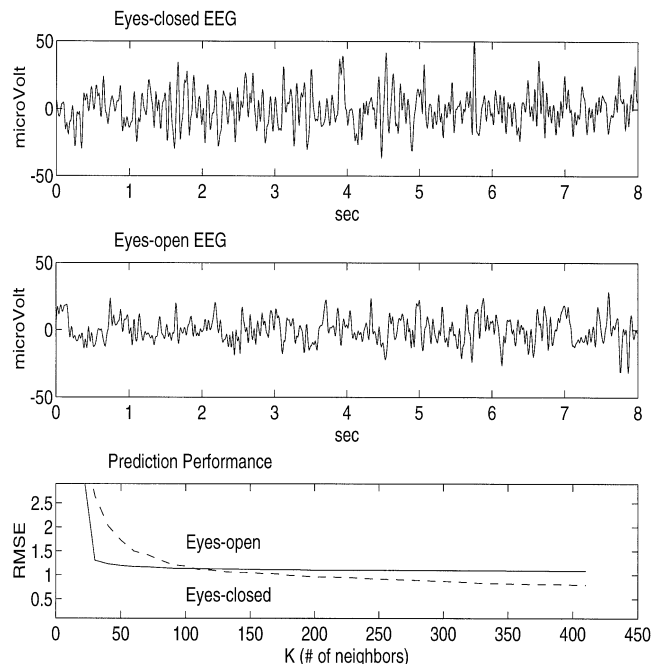


Fig. 4. Prediction of EEG for a subject (#5) having the highest eyes-closed RMSE (dashed line is eyes-open prediction).

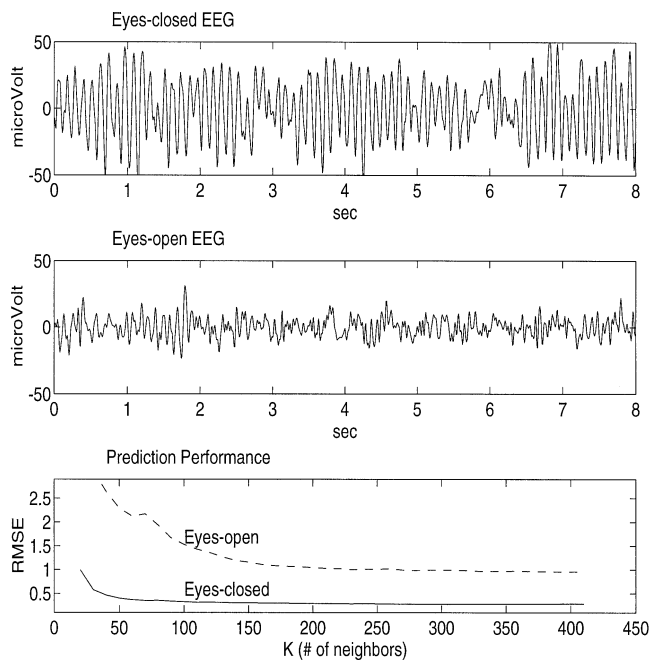


Fig. 5. Prediction of EEG for a subject (#2) having the lowest eyes-closed RMSE (dashed line is eyes-open prediction).

listed in Table 3 it might appear unlikely that the EEGs were produced by a high-dimensional system (either linear stochastic or nonlinear). This conclusion cannot be drawn however, since D_2 estimation

Table 3. Phillips–Perron (P–P) statistic (1% critical value = -3.440), D_2 and ΔD_2 for EEG series (1024 points each).

Subject No.	Eyes-Closed			Eyes-Open		
	P–P	D_2	δD_2	P–P	D_2	δD_2
1	-9.5942	3.3713	0.0976	-9.5548	3.6748	0.0894
	-9.4243	3.3631	0.0677	-12.0971	4.1106	0.0870
	-9.6594	3.4625	0.0260	-9.0619	3.3495	0.1337
2	-11.0411	2.8449	0.2432	-7.6435	3.0097	0.2052
	-10.9801	2.5032	0.3123	-9.8922	3.7230	0.0659
	-10.8120	2.6313	0.2024	-9.5342	3.5383	0.1058
3	-10.5430	2.9872	0.1529	-6.3347	2.7251	0.1511
	-10.4283	2.8921	0.1442	-7.4078	2.9596	0.1829
	-10.8353	2.7157	0.2397	-6.7517	2.8244	0.0989
4	-10.3583	2.8521	0.1982	-7.8824	3.2338	0.1369
	-10.4561	2.8673	0.1184	-6.9780	3.1046	0.2051
	-10.8502	2.6373	0.3548	-6.2722	3.2092	0.0885
5	-8.6500	3.4337	-0.0174	-8.3939	2.9111	0.2719
	-9.5419	3.3867	-0.0318	-7.9507	3.0055	0.1866
	-9.0541	3.1914	0.1121	-7.7311	3.2442	-0.0247
6	-10.6827	2.9785	0.2237	-8.4517	3.6209	0.0254
	-11.0883	3.0392	0.1307	-9.6895	3.6171	0.1711
	-10.5155	3.0275	0.1151	-9.7782	3.3799	0.1158
7	-10.7316	2.8745	0.0936	-7.9589	3.1596	0.0347
	-10.8254	3.0767	0.0440	-9.2705	3.2948	-0.0468
	-10.8885	3.2514	0.0187	-9.3997	3.3615	0.1091
8	-9.4365	2.7425	0.1217	-6.7782	2.7976	0.2061
	-9.4720	2.7402	0.0542	-6.6264	2.8888	0.1970
	-10.0764	2.7051	0.2052	-6.0762	2.7716	0.3079
9	-11.1130	2.8074	0.2944	-8.0106	3.6840	-0.1196
	-10.3420	3.0108	0.0682	-7.8945	3.4555	0.0251
	-11.0925	2.8271	0.1589	-8.9938	3.5415	0.0482
10	-10.5869	2.5712	0.3823	-7.6819	3.5259	0.0102
	-10.8914	2.9416	0.1428	-8.1037	3.1506	0.1858
	-10.2390	2.9556	0.0438	-8.4186	3.3486	0.0633
Mean	-10.3403	2.9563	0.1439	-8.2206	3.2740	0.1106

algorithms are often “fooled” by stochastic systems having $1/f$ -like power spectra [Theiler, 1991] such as the EEG. Theiler [1991] pointed out that “The concept of fractal dimension can be applied to time series in two quite distinct ways: one, to indicate the number of degrees of freedom in the underlying dynamical system; and two, to quantify the self-affinity (or ‘crinkliness’) of the trajectory through

phase space.” It is in this latter sense that we employed D_2 as a general index of dynamical complexity, thereby making no inferences regarding the “true dimension” or the absolute degrees of freedom of the EEG.

The results of our study support the hypothesis that these EEGs were produced by a linear stochastic system. They concretely demonstrate that EEG

prediction is carried out sufficiently using linear AR modeling, and that nonlinear prediction (such as by the KNNLL method) does not explain any additional sources of variation in the data. This conclusion corroborates a similar one drawn in a recent study based on the analysis of EEGs using correlation dimension [Theiler & Rapp, 1996]. In that particular study (which was a reanalysis of a 1989 study), Theiler and Rapp [1996] concluded that with respect to EEG dimension estimation “greater success may be obtained with measures that are sensitive to other aspects of a system’s structure and which may have lower data requirements. Candidate measures include predictability . . .” Our results here tend to corroborate this opinion with the caveat that they apply to a limited normative data set. However, the combination of prediction analysis along with simulation studies using time series with well-described dynamics impels us to conclude that the probability that this particular data set arises from a nonlinear system is quite low.

Hernandez *et al.* [1995] used similar prediction methodology on 26 eyes-closed EEG segments of normals and patients and concluded that in 18 cases the data was better fit using the nonlinear SETAR (self-exciting threshold AR) model attributed to Tong [1990]. Comparison of these results to our study is difficult since we do not know from which scalp recording sites Hernandez *et al.* obtained their EEG data from, nor which data was from patients. They conceded that noise was the main contributor in low predictability of alpha EEG activity. Our approach differs in that direct comparison of results between simulated time series of known dimensional complexity in the presence of varying noise levels with both eyes-open and eyes-closed EEGs was carried out in a repeated measure analysis.

The finding that, for the eyes-closed data, greater stationarity was associated with a less complex reconstructed state-space trajectory (i.e. lower estimated D_2) makes intuitive sense, since in the hypothetical limit of “perfect stationarity”, one can imagine the alpha rhythm approaching a sine wave (dimension 1 and perfectly predictable). Less intuitively obvious is why greater stationarity should be associated with *higher* estimated D_2 in the eyes-open control data. Perhaps one may speculate that in the hypothetical limit of “perfect stationarity”, desynchronized EEG would approach a white noise-like state of infinite dimension.

Somewhat puzzling is our finding that δD_2 correlated with both linear and nonlinear predictability. If δD_2 indexes dynamical nonlinearity, as it is supposed to, then there should be no correlation between it and any measure of predictability of a linear process (again, we found no evidence that the alpha rhythm arises from a nonlinear system). δD_2 could also represent “nonlinearity” due, not to a dynamic but to a static nonlinear measurement transform of a linear, non-Gaussian process. This represents the fact that the normal surrogate data procedure tends to “whiten” the spectrum of the surrogates for highly periodic data such as alpha (which increases D_2 estimates) or may represent some type of nonstationarity in the data not reflected by the P–P statistic. This appears to indicate limits on the validity of the phase-angle-randomization surrogate data technique, which suffers from the so-called “edge effects” [Theiler & Rapp, 1996] as well as limited FFT frequency resolution when applied to relatively short data sets such as ours (8 s, yielding an FFT frequency resolution of 0.125 Hz); for an allusion to small sample sizes producing problems for phase-angle-randomized surrogates see [Theiler & Prichard, 1996].

References

- Babloyantz, A., Salazar, J. M. & Nicolis, J. [1985] “Evidence of chaotic dynamics of brain activity during the sleep cycle,” *Phys. Lett.* **A111**, 152–156.
- Blinowska, K. J. & Malinowski, M. [1991] “Nonlinear and linear forecasting of the EEG time series,” *Biol. Cybern.* **66**, 159–165.
- Cao, L., Hong, Y., Fang, H. & He, G. [1995] “Predicting chaotic time series with wavelet networks,” *Physica* **D85**, 2238–2255.
- Casdagli, M. [1989] “Nonlinear prediction of chaotic time series,” *Physica* **D35**, 335–356.
- Casdagli, M. [1992] “Chaos and deterministic versus stochastic nonlinear modeling,” *J. R. Stat. Soc.* **B54**, 303–328.
- Ellner, S. [1988] “Estimating attractor dimensions from limited data: A new method, with error estimate,” *Phys. Lett.* **A133**, 128–133.
- Farmer, J. D. & Sidorowich, J. J. [1988] “Exploiting chaos to predict the future and reduce noise,” Los Alamos, LA-UR-88-901.
- Grassberger, P. & Procaccia, I. [1983] “Measuring the strangeness of strange attractors,” *Physica* **D9**, 189–208.
- Hernandez, J. L., Valdes, J. L., Biscay, R., Jimenez,

- J. C. & Valdes, P. [1995] "EEG predictability: Adequacy of non-linear forecasting methods," *Int. J. Biomed. Comput.* **38**, 107–206.
- Iasemidis, L. D., Sackellares, J. C., Zaveri, H. P. & Williams, W. J. [1990] "Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures," *Brain Topog.* **2**, 187–201.
- Jansen, H. [1991] "Quantitative analysis of electroencephalograms: Is there chaos in the future?" *Int. J. Biomed. Comput.* **27**, 95–123.
- Kemp, B. & Lopes da Silva, F. H. [1991] "Model based analysis of neurophysiological signals," *Digital Biosignal Processing*, ed. Weikunat, R. (Elsevier, Amsterdam).
- Lowe, D. & Webb, A. R. [1991] "Time series prediction by adaptive networks: A dynamical systems perspective," *IEE Proc.* **F138**, 17–24.
- Mann, N. R., Schafer, R. E. & Singpurwalla, N. D. [1974] *Methods for Statistical Analysis of Reliability and Life Data* (Wiley, NY).
- Mayer-Kress, G. & Layne, S. P. [1987] "Dimensionality of the human electroencephalogram," *Ann N.Y. Acad. Sci.* **504**, 62–86.
- Molinari, L. & Dumermuth, G. [1992] "Once more: Dimension and Lyapunov exponents for the human EEG," *Nonlinear Dynamical Analysis of the EEG*, eds. Jansen, B. H. & Brandt, M. E. (World Scientific, Singapore), pp. 140–155.
- Newey, W. & West, K. [1987] "A simple positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica* **55**, 703–708.
- Palus, M. [1993] "Testing for nonlinearity in the EEG," *Nonlinear Dynamical Analysis of the EEG*, eds. Jansen, B. H. & Brandt, M. E. (World Scientific, Singapore), pp. 100–114.
- Palus, M., Albrecht, V. & Dvorak, I. [1993] "Information theoretic test for nonlinearity in time series," *Phys. Lett.* **A175**, 203–209.
- Parikh, J. C. & Pratap, R. [1991] "A map describing EEG activity of human brain," *Pramana J. Phys.* **36**, L347–L352.
- Pezard, L., Martinerie, J., Breton, F., Bourzeix, J. & Renault, B. [1994] "Nonlinear forecasting measurements of multichannel EEG dynamics," *Electroenceph. Clin. Neurophysiol.* **91**, 383–391.
- Phillips, P. C. B. & Perron, P. [1988] "Testing for a unit root in time series regression," *Biometrika* **75**, 335–346.
- Principe, J. C. & Lo, P. [1991] "Towards the determination of the largest Lyapunov exponent of EEG segments," *Measuring Chaos in the Human Brain*, eds. Duke, D. W. & Pritchard, W. S. (World Scientific, Singapore), pp. 156–166.
- Pritchard, W. S. & Duke, D. W. [1992a] "Dimensional analysis of no-task human EEG using the Grassberger-Procaccia method," *Psychophysiology* **29**, 182–192.
- Pritchard, W. S. & Duke, D. W. [1992b] "Measuring chaos in the brain: A tutorial review of nonlinear dynamical EEG analysis," *Int. J. Neurosci.* **67**, 31–80.
- Pritchard, W. S., Duke, D. W. & Kriehle, K. K. [1995] "Dimensional analysis of resting human EEG II: Surrogate-data testing indicates nonlinearity but not low-dimensional chaos," *Psychophysiology* **32**, 486–491.
- Rapp, P. E., Bashore, T. R., Martinerie, J. M., Albano, A. M. & Mees, A. I. [1989] "Dynamics of brain electrical activity," *Brain Topog.* **2**, 99–118.
- Röschke, J. & Başar, E. [1989] "Correlation dimension in various parts of cat and human brain in different states," *Brain Dynamics, Progress and Perspectives*, Vol. 2, eds. Başar, E. & Bullock, T. H. (Springer, Berlin), pp. 131–148.
- Rosenstein, M. T., Collins, J. J. & De Luca, C. J. [1994] "Reconstruction expansion as a geometry-based framework for choosing proper delay times," *Physica* **D73**, 82–98.
- Stam, C. J., Jelles, B., Achtereekte, H. A. M., Rombouts, S. A. R. B., Slaets, J. P. J. & Keunen, R. W. M. [1995] "Investigation of EEG non-linearity in dementia and Parkinson's disease," *Electroenceph. Clin. Neurophys.* **95**, 309–317.
- Stokbro, K. & Umberger, D. K. [1992] "Forecasting with weighted maps," *Nonlinear Modeling and Forecasting, SFI Studies in the Sciences of Complexity*, Vol. 12, eds. Casdagli, M. & Eubank, S. (Addison-Wesley, Reading), pp. 73–93.
- Sugihara, G. & May, R. M. [1990] "Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series," *Nature* **344**, 734–741.
- Takens, F. [1981] "Detecting strange attractors in turbulence," *Lecture Notes in Mathematics* **898**, 366–381.
- Takens, F. [1985] "On the numerical determination of the dimension of an attractor," *Lecture Notes in Mathematics* **1125**, 99–115.
- Theiler, J. [1991] "Some comments on the correlation dimension of $1/f^\alpha$ noise," *Phys. Lett.* **A155**, 480–493.
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Farmer, J. D. [1992] "Testing for nonlinearity in time series: The method of surrogate data," *Physica* **D58**, 77–94.
- Theiler, J. & Prichard, D. [1996] "Using 'surrogate data' to calibrate the actual rate of false positives in tests for nonlinearity in time series," *Fields Inst. Commun.* **11**, 99–111.
- Theiler, J. & Rapp, P. E. [1996] "Re-examination of the evidence for low-dimensional, nonlinear structure

- in the human electroencephalogram,” *Electroenceph. Clin. Neurophysiol.* **98**, 213–222.
- Tong, H. [1990] *Nonlinear Time Series Analysis: A Dynamical Systems Approach* (Oxford University Press, Oxford).
- Wallenstein, G. V. & Nash, A. J. [1992] “Does the largest Lyapunov characteristic exponent quantify the dynamical complexity of the event-related potential? An algorithm using adaptive step-size control for variable propagation times,” *Nonlinear Dynamical Analysis of the EEG*, eds. Jansen, B. H. & Brandt, M. E. (World Scientific, Singapore), pp. 232–256.
- Weigend, A. S., Huberman, B. A. & Rumelhart, D. E. [1992] “Predicting sunspots and exchange rates with connectionist networks,” *Nonlinear Modeling and Forecasting, SFI Studies in the Sciences of Complexity*, Vol. 12, eds. Casdagli, M. & Eubank, S. (Addison-Wesley, Reading), pp. 395–432.